NUWC-NPT Technical Report 11,711
30 September 2005

Second-Order Statistics of Spectral and Correlation Estimates Obtained by Means of Weighted Overlapped FFT Processing

Albert H. Nuttall
Sensors and Sonar Systems Department

NAVSEA WARFARE CENTERS
NEWPORT

Naval Undersea Warfare Center Division
Newport, Rhode Island

PREFACE

The work described in this report was funded under the Sonar Automatic Technology project, principal investigators Stephen G. Greineder (Code 159B) and Mary H. Johnson (Code 1522). The sponsoring activity is the Office of Naval Research, program managers J. Tague and H. South (ONR 321US).

The technical reviewer for this report was Paul M. Baggenstoss (Code 1522).

Reviewed and Approved: 30 September 2005

Joseph M. Monti
Head, Sensors and Sonar Systems Department

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	30 September 2005	

4. TITLE AND SUBTITLE

Second-Order Statistics of Spectral and Correlation Estimates Obtained by Means of Weighted Overlapped FFT Processing

5. FUNDING NUMBERS

PR A101576

6. AUTHOR(S)

Albert H. Nuttall

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Undersea Warfare Center Division
1176 Howell Street
Newport, RI 02841-1708

8. PERFORMING ORGANIZATION REPORT NUMBER

TR 11,711

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research (ONR 321MS)
Ballston Centre Tower One
800 North Quincy Street
Arlington VA 22217-5660

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Closed-form relations have been obtained for the covariances of the spectral estimates, as well as the correlation estimates, that result from weighted overlapped FFT processing of a stationary process with Gaussian characteristics. In particular, estimates that result from time-displaced overlapped data segments and different frequency bins have been considered. The effects of wraparound in the time-delay domain of the correlation estimates have also been incorporated. The fundamental parameters of the study are the weighting segment length K, the amount of lag L, and the FFT size N, as well as the input data correlation function $C(k)$ and the particular weighting sequence $w(k)$.

14. SUBJECT TERMS

Correlation Estimation	Spectral Estimation
Second-Order Statistics	Weighted FFT
Signal Processing	

15. NUMBER OF PAGES

46

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

TABLE OF CONTENTS

Section **Page**

LIST OF ABBREVIATIONS, ACRONYMS, AND SYMBOLS ii

1 INTRODUCTION 1

2 STATISTICS OF MAGNITUDE-SQUARED SPECTRAL ESTIMATES 3
Non-White Gaussian Input Data 6
Relation to Correlation Estimates of the Input Data 10

3 STATISTICS OF CORRELATION ESTIMATES FOR WHITE DATA AND N ≥ K 13

4 STATISTICS OF CORRELATION ESTIMATES FOR NON-WHITE DATA AND N ≥ 2K 19

5 STATISTICS OF CORRELATION ESTIMATES FOR NON-WHITE DATA AND N ≥ K 23

6 STATISTICS OF COMPLEX SPECTRAL ESTIMATES 27

7 SUMMARY 31

APPENDIX A—INVERSE DISCRETE FOURIER TRANSFORM PROPERTIES A-1

APPENDIX B—APPROXIMATE EVALUATION OF $\theta(m,n,p)$ B-1

LIST OF ILLUSTRATIONS

Figure **Page**

1 MATLAB Example for the Covariance Matrix 7
2 MATLAB Example for Equation (91) 18

LIST OF ABBREVIATIONS, ACRONYMS, AND SYMBOLS

A	Auxiliary ensemble average, equation (73)
$A(\underline{n}, n+\underline{n})$	Auxiliary window function, equation (143)
B	Auxiliary ensemble average, equation (74)
$B(\underline{n}, n-\underline{n})$	Auxiliary ensemble average, equation (144)
CAF	Complex ambiguity function, equation (14)
CC	Crosscorrelation, equations (10) and (72)
$C(k)$	Correlation of input data, equation (26)
$C_N(k)$	Inverse discrete Fourier transform, equation (A-2)
cov	Covariance, equation (18)
C_{qq}	Covariance matrix, equation (50)
C_{qp}	Covariance matrix, equation (52)
$e(m)$	Auxiliary vector, equation (46)
E	Expectation, equation (7)
$f(m)$	Auxiliary vector, equation (54)
FFT	Fast Fourier transform
$F(L, n)$	Auxiliary function, equation (22)
$G(k)$	Gate function, equations (30) and (A-6)
$\mathbf{g}(k)$	Gaussian random input sample at sample time k, equation (1)
$h(k, m)$	Product of weighting sequences, equations (81) and (130)
$H(m, L)$	Autocorrelation of $h(k, L)$, equation (83)
$\underline{H}(p, m, \underline{m})$	Autocorrelation of $h(k, m)$, equation (129)
$k, \underline{k}$	Time indexes, equations (1) and (10)
K	Duration of weighting sequence, equation (1)
L	Lag of weighting sequence, equation (3)
$m, \underline{m}$	Time-delay indices, equation (98)
mod	Modulo function, equation (24)
M	Integer $N/2+1$, below equation (45)
$n, \underline{n}$	Frequency bin numbers, equations (1) and (6)
N	Size of FFT, equation (1)
$\mathbf{p}(k)$	Product sequence, equation (95)
$\mathbf{p}(n)$	Magnitude-squared spectral estimate of lagged data, equation (6)
P, Q, R, S	Four auxiliary averages, equation (112)
$\mathbf{q}(m)$	Correlation estimate, equation (62)
$\mathbf{q}(n)$	Magnitude-squared spectral estimate, equation (5)
$\mathbf{r}(m)$	Correlation estimate, equations (42), (61), and (92)

LIST OF ABBREVIATIONS, ACRONYMS, AND SYMBOLS (Cont'd)

RV	Random variable
$\mathbf{s}(m)$	Correlation estimate, equation (51)
$S(m)$	Spectrum of input data, equation (30)
$S_N(m)$	Discrete Fourier transform, equation (A-1)
Sub i	Imaginary part of function, equation (142)
Sub r	Real part of function, equation (142)
$T(L, m, \underline{m})$	Combination function, equation (124)
$\mathbf{u}(n)$	Real part of $\tilde{\mathbf{z}}(\underline{n})$, equation (145)
$\mathbf{v}(n)$	Imaginary part of $\tilde{\mathbf{z}}(\underline{n})$, equation (145)
var	Variance, equation (20)
$w(k)$	Weighting sequence at sample time k, equation (1)
$W(n)$	Window corresponding to weighting sequence, equation (31)
$W_1(n)$	Window for lagged sequence product, equation (138)
$W_2(n)$	Window for squared weighting sequence, equation (139)
$\mathbf{x}(n)$	Real part of $\mathbf{z}(n)$, above equation (142)
$\mathbf{y}(n)$	Imaginary part of $\mathbf{z}(n)$, above equation (142)
$\mathbf{y}(n)$	Complex spectral estimate of lagged data, equation (3)
$\mathbf{z}(n)$	Complex spectral estimate, equation (1)
$\tilde{\mathbf{z}}(n)$	Complex spectral estimate for lagged data, equation (137)
$\delta(k)$	Kronecker delta function, equation (8)
$\delta_N(m)$	Periodic Kronecker delta function, equation (70)
$\phi_w(m)$	Autocorrelation of weight sequence, equation (29)
$\theta(m, n, p)$	Third-order autocorrelation of weighting sequence, equation (104)

SECOND-ORDER STATISTICS OF SPECTRAL AND CORRELATION ESTIMATES OBTAINED BY MEANS OF WEIGHTED OVERLAPPED FFT PROCESSING

1. INTRODUCTION

An efficient method of estimating the spectral and/or correlation characteristics of a stationary random digital temporal process, $\{\mathbf{g}(k)\}$, is by the use of weighted overlapped fast Fourier transform (FFT) procedures. In particular, a weighting sequence $\{w(k)\}$ of finite duration K is overlaid on the long input data sequence $\{\mathbf{g}(k)\}$ and the product is subjected to an FFT of size N, yielding complex frequency coefficients. Parameter N governs the frequency spacing of the coefficients. The magnitude-squared frequency coefficients for this weighted segment of data are stored in computer memory. Then, the weighting sequence is lagged (delayed) by L units of time and the FFT procedure is repeated on the weighted input data. When a sufficient number of these frequency-transformed segments are available, each with a different time lag L, an average is formed of the magnitude-squared frequency coefficients, thereby yielding an estimate of the power density spectrum of the input data. If and when an estimate of the correlation function of the input data is also desired, an inverse FFT of the magnitude-squared frequency coefficients is performed, yielding correlation estimates in the time-delay domain.

The joint higher-order probability density functions of the spectral and/or correlation estimates are often of importance in signal processing. However, there are features of the FFT processing above that make this virtually impossible analytically. Even if input data $\{\mathbf{g}(k)\}$ are zero-mean, white, and Gaussian, the overlap of adjacent weighting sequences, that is, lag L less than segment duration K, coupled with the nonlinear operation of magnitude-squared frequency coefficients and the subsequent averaging, leads to statistical problems that are analytically intractable. Although the joint moment-generating function can often be calculated for some of these types of processing operations, the required matrix manipulations have storage and size requirements that are frequently not practical. Also, the interaction of the fundamental parameters, namely, weighting duration K, temporal lag L, and FFT size N, lead to complications in the analysis.

Accordingly, the approach here is limited to determination of the first-order and second-order statistics of the spectral and correlation estimates. Several cases are considered, including both white and colored Gaussian input data. Lag L is arbitrary, thereby allowing for study of the adjacent time segments, as well as the separated (disjoint) time segments of data. The FFT size N is also arbitrary, thereby allowing for the effects of wraparound on the estimates and their stability. In particular, the dependence between estimates obtained at different segment locations and at different frequency bins is investigated, and closed-form results for the means and covariances are obtained.

Section 2 of this report contains the derivation of the second-order statistics of the magnitude-squared spectral estimates. Then, three different cases of correlation estimates are investigated in sections 3, 4, and 5, each progressively more general and more difficult analytically. Section 6 treats the statistics of the *complex* frequency coefficients themselves. Finally, appendixes A and B present detailed investigations into several special topics.

2. STATISTICS OF MAGNITUDE-SQUARED SPECTRAL ESTIMATES

Input data sequence $\{\mathbf{g}(k)\}$ consists of independent, real, Gaussian, random variables (RVs) with zero mean and unit variance. Real weighting sequence $\{w(k)\}$ is defined for all integer k, but is nonzero only for $k = 1:K$. A complex spectral estimate in frequency bin n is obtained according to

$$\mathbf{z}(n) = \sum_k \mathbf{g}(k)\, w(k) \exp(-i2\pi\, n\, k/N) \quad \text{for } n = 0:N-1. \tag{1}$$

The infinite sum on k is automatically terminated to K terms by the weighting; this isolates a segment of the data stream $\{\mathbf{g}(k)\}$. The integer N governs the frequency spacing of the complex spectral estimates $\{\mathbf{z}(n)\}$. Since

$$\mathbf{z}(N-n) = \mathbf{z}^*(n) \quad \text{for } n = 1:N-1, \tag{2}$$

it is only necessary to consider the range $n = 0:N/2$ for $\{\mathbf{z}(n)\}$, where N is presumed even.

In addition, another set of complex spectral estimates is obtained from a lagged data set relative to sequence $\{\mathbf{g}(k)\}$, according to

$$\mathbf{y}(n) = \sum_k \mathbf{g}(k)\, w(k-L) \exp[-i2\pi\, n\,(k-L)/N] = \sum_k \mathbf{g}(k+L)\, w(k) \exp(-i2\pi\, nk/N), \tag{3}$$

where L is the amount of lag between the two sets of data and is arbitrary. If lag L is greater than or equal to weighting duration K, the two sets of spectral estimates in equations (1) and (3) are statistically independent of each other, because they involve different sets of independent, Gaussian RVs. Interest here will be concentrated on the case $0 \le L < K$, meaning that the two segmented data sets encountered in equations (1) and (3) are statistically dependent on each other, thereby making complex RVs $\{\mathbf{z}(n)\}$ and $\{\mathbf{y}(n)\}$ statistically dependent on each other. It is also assumed that $N \ge K$; thus, the various integers of the following spectral analysis will always satisfy

$$0 \le L < K \le N; \tag{4}$$

however, L can be arbitrary in general.

The magnitude-squared spectral estimates of interest are

$$\mathbf{q}(n) = |\mathbf{z}(n)|^2 = \sum_{k,j} \mathbf{g}(k)\, \mathbf{g}(j)\, w(k)\, w(j) \exp[-i2\pi\, n\,(k-j)/N] \quad \text{for } n = 0:N/2, \tag{5}$$

and

$$\mathbf{p}(\underline{n}) = |\mathbf{y}(\underline{n})|^2 = \sum_{k,j} \mathbf{g}(k+L)\,\mathbf{g}(j+L)\; w(k)\; w(j) \exp[-i2\pi \underline{n}(k-j)/N] \;\text{ for } \underline{n} = 0 : N/2, \tag{6}$$

where integer $\underline{n}$ need not be equal to n. The autocorrelation of real data sequence $\{\mathbf{g}(k)\}$ is

$$E\{\mathbf{g}(k)\,\mathbf{g}(j)\} = \delta(k-j), \tag{7}$$

where $E\{\,\}$ denotes an ensemble average, and the Kronecker delta is defined here as

$$\delta(k) = \begin{Bmatrix} 1 \text{ for } k=0 \\ 0 \text{ otherwise} \end{Bmatrix}. \tag{8}$$

Then, from equation (5), the mean of magnitude-squared spectral estimate $\mathbf{q}(n)$ is immediately

$$E\{\mathbf{q}(n)\} = \sum_k w^2(k) \quad \text{for all } n. \tag{9}$$

Of course, spectral estimates $\{\mathbf{p}(\underline{n})\}$ have the same common mean value.

The crosscorrelation of the two spectral estimates in equations (5) and (6), for two different (or equal) frequency indices n and $\underline{n}$, is

$$\begin{aligned} CC \equiv E\{\mathbf{q}(n)\,\mathbf{p}(\underline{n})\} = \sum_{kj\underline{k}\underline{j}} E\{\mathbf{g}(k)\,\mathbf{g}(j)\,\mathbf{g}(\underline{k}+L)\,\mathbf{g}(\underline{j}+L)\}\; w(k)\; w(j)\; w(\underline{k})\; w(\underline{j}) \\ \times \exp[-i2\pi\{n(k-j)+\underline{n}(\underline{k}-\underline{j})\}/N] \;\text{ for } n=0:N/2, \underline{n}=0:N/2. \end{aligned} \tag{10}$$

Since data sequence $\{\mathbf{g}(k)\}$ is Gaussian, the ensemble average in equation (10) can be written, with the help of equation (7), as a sum of three terms, namely,

$$\delta(k-j)\,\delta(\underline{k}-\underline{j}) + \delta(k-\underline{k}-L)\,\delta(j-\underline{j}-L) + \delta(k-\underline{j}-L)\,\delta(j-\underline{k}-L). \tag{11}$$

Substitution of the first product of Kronecker delta functions into equation (10) yields the first component of the crosscorrelation CC as

$$CC_1 = \sum_{k,\underline{k}} w^2(k)\, w^2(\underline{k}) = \left(\sum_k w^2(k)\right)^2 = E\{\mathbf{q}(n)\}\, E\{\mathbf{p}(\underline{n})\}. \tag{12}$$

Thus, the sum of the two remaining components of CC will be the covariance of RVs $\mathbf{q}(n)$ and $\mathbf{p}(\underline{n})$.

Substitution of the second product of equation (11) into equation (10) yields the second component of CC as

$$CC_2 = \sum_{k,j} w(k)\, w(k-L)\, w(j)\, w(j-L)\, \exp[-i2\pi(n+\underline{n})(k-j)/N]$$
$$= \left| \sum_{k} w(k)\, w(k-L)\, \exp[-i2\pi(n+\underline{n})\,k/N] \right|^2 . \tag{13}$$

At this point, it is useful to define the complex ambiguity function (CAF) of real weighting sequence $\{w(k)\}$ as

$$\chi(m,n) = \sum_{k} w(k)\, w(k-m)\, \exp(-i2\pi\, n\, k/N) \text{ for all } m, n. \tag{14}$$

Then,

$$CC_2 = \left|\chi(L, n+\underline{n})\right|^2 . \tag{15}$$

The CAF $\chi(m,n)$ has period N in variable n. Also,

$$\chi(L,n) = 0 \text{ if } L \ge K, \quad \chi(L, N-n) = \chi^*(L,n). \tag{16}$$

The third component of CC is obtained by substituting the third product of equation (11) into equation (10), namely,

$$CC_3 = \sum_{k,j} w(k)\, w(k-L)\, w(j)\, w(j-L)\, \exp[-i2\pi(n-\underline{n})(k-j)/N] = \left|\chi(L, n-\underline{n})\right|^2 . \tag{17}$$

Thus, as noted under equation (12), the covariance of interest is

$$\operatorname{cov}\{\mathbf{q}(n), \mathbf{p}(\underline{n})\} = \left|\chi(L, n-\underline{n})\right|^2 + \left|\chi(L, n+\underline{n})\right|^2 \text{ for } n = 0:N/2,\ \underline{n} = 0:N/2. \tag{18}$$

Equations (18) and (14) are the main results. Lag L is arbitrary; however, if L is taken larger than or equal to K, the CAF in equation (14) will be zero, meaning that all the spectral estimates $\{\mathbf{q}(n)\}$ and $\{\mathbf{p}(\underline{n})\}$ are uncorrelated with each other for all $n, \underline{n}$. As a special case of equation (18), by setting $L = 0$, there follows

$$\operatorname{cov}\{\mathbf{q}(n), \mathbf{q}(\underline{n})\} = \left|\chi(0, n-\underline{n})\right|^2 + \left|\chi(0, n+\underline{n})\right|^2 \text{ for } n = 0:N/2,\ \underline{n} = 0:N/2, \tag{19}$$

which is the covariance of just the spectral estimates $\{\mathbf{q}(n)\}$. As a further special case,

$$\operatorname{var}\{\mathbf{q}(n)\} = \chi(0,0)^2 + \left|\chi(0,2n)\right|^2 \text{ for } n = 0:N/2. \tag{20}$$

Thus, for example,

$$\mathrm{var}\{\mathbf{q}(0)\} = \mathrm{var}\{\mathbf{q}(N/2)\} = 2\ \chi(0,0)^2. \tag{21}$$

For $0 \le L < K$, define

$$F(L,n) = \sum_k w(k)\ w(k-L)\ \exp(-i2\pi\, n\, k/N) \quad \text{for } n = 0 : N-1. \tag{22}$$

This quantity can be realized in MATLAB as an FFT:

$$F(L,:) = \mathrm{fft}(w(L+1:K).*\,w(1:K-L), N). \tag{23}$$

Then, the CAF is available according to

$$\chi(L,n) = F(L,1+\mathrm{mod}(n,N)) \quad \text{for all } n. \tag{24}$$

Since integers n and $\underline{n}$ are limited to $0 : N/2$, the mod function comes into play when

$$n = N/2 \ \text{ and } \ \underline{n} = N/2, \quad \text{or} \quad n + \underline{n} = N. \tag{25}$$

Finally, the mean of the spectral estimates in equation (9) can be expressed in terms of the CAF as $E\{\mathbf{q}(n)\} = \chi(0,0)$ for all n. An example MATLAB program for the covariance matrix is shown in figure 1.

NON-WHITE GAUSSIAN INPUT DATA

If the input data correlation in equation (7) is replaced by

$$E\{\mathbf{g}(k)\ \mathbf{g}(j)\} = C(k-j), \quad C(0) = 1, \tag{26}$$

which is a real even function, then the mean in equation (9) is replaced by

$$E\{\mathbf{q}(n)\} = \sum_{k,j} C(k-j)\ w(k)\ w(j)\ \exp[-i2\pi\, n(k-j)/N]. \tag{27}$$

Letting $m = k - j$, there follows

$$E\{\mathbf{q}(n)\} = \sum_{k,m} C(m)\ w(k)\ w(k-m)\ \exp(-i2\pi\, nm/N) = \sum_m \exp(-i2\pi\, nm/N)\ C(m)\ \phi_w(m), \tag{28}$$

where

```
clear all
close all hidden

K=1024;                 % Length of weighting
L=512;                  % Lag (shift) of adjacent segment; 0
<= L < K
N=2048;                 % FFT size; K <= N
w=hamming(K);           % Weighting sequence
% w=hann(K);
% w=hann(K+1); w=w(2:K+1);
% w=hann(K+2); w=w(2:K+1);

average=sum(w.^2);
F=fft(w(L+1:K).*w(1:K-L),N);     % F Depends on lag L
F2=F.*conj(F);
cov2=zeros(N/2+1,N/2+1);
for na=0:N/2
    for nb=0:N/2
        td=F2(1+mod(na-nb,N));
        ts=F2(1+mod(na+nb,N));
        cov2(na+1,nb+1)=td+ts;
    end
end
```

Figure 1. MATLAB Example for the Covariance Matrix

$$\phi_w(m) = \sum_k w(k)\, w(k-m) \tag{29}$$

is the autocorrelation of weight sequence $\{w(k)\}$, and $\phi_w(m) = 0$ for $|m| \geq K$.

On the other hand, if $N \geq 2K$, and if the input data correlation $C(k)$ is expressed in terms of the real and even input data spectrum $S(m)$ according to (see appendix A)

$$C(k) = G(k) \sum_m \exp(i2\pi\, km/N)\, S(m), \quad G(k) = \begin{Bmatrix} 1 & \text{for } |k| < N/2 \\ 0 & \text{for } |k| \geq N/2 \end{Bmatrix}, \tag{30}$$

and if window $W(n)$ is defined as

$$W(n) = \sum_k \exp(-i2\pi\, nk/N)\, w(k), \tag{31}$$

then substitution into equation (28) leads to mean value

$$E\{\mathbf{q}(n)\} = \sum_m S(m)\, |W(n-m)|^2. \tag{32}$$

The product $G(m)\,\phi_w(m)$ that occurs upon substitution into equation (28) is equal to $\phi_w(m)$ for all m when $N \ge 2K$. On the other hand, if only the weaker restriction $N \ge K$ is in effect, the gating function $G(m)$ partially truncates $\phi_w(m)$, thereby obviating the simple relation (32).

Relation (32) is recognized as a frequency domain convolution of the input data spectrum with the magnitude-squared window of the weighting sequence. Equation (28) is a correlation domain expression for the mean of the magnitude-squared spectral estimate $\mathbf{q}(n)$, whereas equation (32) is a frequency domain expression for this same quantity. However, relation (32) requires that $N \ge 2K$, whereas relation (28) only requires $N \ge K$.

For non-white Gaussian input data, the crosscorrelation CC of two different spectral estimates $\mathbf{q}(n)$ and $\mathbf{p}(\underline{n})$ is still given by general relation (10). However, equation (11) is now replaced by

$$\begin{aligned} &E\{\mathbf{g}(k)\,\mathbf{g}(j)\,\mathbf{g}(\underline{k}+L)\,\mathbf{g}(\underline{j}+L)\} = C(k-j)\,C(\underline{k}-\underline{j}) \\ &\quad + C(k-\underline{k}-L)\,C(j-\underline{j}-L) + C(k-\underline{j}-L)\,C(j-\underline{k}-L). \end{aligned} \tag{33}$$

Substitution of the first product of equation (33) into equation (10) yields the first component of the crosscorrelation CC as

$$\begin{aligned} CC_1 &= \sum_{kj\underline{k}\underline{j}} C(k-j)\,C(\underline{k}-\underline{j})\,w(k)\,w(j)\,w(\underline{k})\,w(\underline{j})\exp[-i2\pi\{n(k-j)+\underline{n}(\underline{k}-\underline{j})\}/N] \\ &= \sum_{k,j} C(k-j)\,w(k)\,w(j)\exp[-i2\pi\, n(k-j)/N] \\ &\quad * \sum_{\underline{k},\underline{j}} C(\underline{k}-\underline{j})\,w(\underline{k})\,w(\underline{j})\exp[-i2\pi\underline{n}(\underline{k}-\underline{j})/N] \\ &= E\{\mathbf{q}(n)\}\,E\{\mathbf{p}(\underline{n})\}. \end{aligned} \tag{34}$$

Thus, the sum of the two remaining components of CC will be the covariance of RVs $\mathbf{q}(n)$ and $\mathbf{p}(\underline{n})$.

Substitution of the second product of equation (33) into equation (10) yields the second component of CC as

$$\begin{aligned} CC_2 &= \sum_{kj\underline{k}\underline{j}} C(k-\underline{k}-L)\,C(j-\underline{j}-L)\,w(k)\,w(j)\,w(\underline{k})\,w(\underline{j})\exp[-i2\pi\{n(k-j)+\underline{n}(\underline{k}-\underline{j})\}/N] \\ &= \sum_{k,\underline{k}} C(k-\underline{k}-L)\,w(k)\,w(\underline{k})\exp[-i2\pi(nk+\underline{n}\underline{k})/N] \\ &\quad * \sum_{j,\underline{j}} C(j-\underline{j}-L)\,w(j)\,w(\underline{j})\exp[+i2\pi(nj+\underline{n}\underline{j})/N]. \end{aligned} \tag{35}$$

In the $k,\underline{k}$ sum, let $\underline{k} = k - m$ to get

$$
\begin{aligned}
&\sum_{k,m} C(m-L)\, w(k)\, w(k-m) \exp[-i2\pi nk/N - i2\pi \underline{n}(k-m)/N] \\
&= \sum_{m} \exp(i2\pi \underline{n}m/N)\, C(m-L) \sum_{k} w(k)\, w(k-m) \exp[-i2\pi(n+\underline{n})k/N] \\
&= \sum_{m} \exp(i2\pi \underline{n}m/N)\, C(m-L)\, \chi(m, n+\underline{n}),
\end{aligned}
\tag{36}
$$

upon use of equation (14). A similar procedure applied to the $j,\underline{j}$ sum in equation (35) leads to

$$
\begin{aligned}
CC_2 &= \left| \sum_{m} \exp(i2\pi \underline{n}m/N)\, C(m-L)\, \chi(m, n+\underline{n}) \right|^2 \\
&= \left| \sum_{p} \exp(i2\pi \underline{n}p/N)\, C(p)\, \chi(L+p, n+\underline{n}) \right|^2 .
\end{aligned}
\tag{37}
$$

The third component of CC is obtained by substituting the third product of equation (33) into equation (10). Then, a procedure similar to that employed in equations (36) and (37) yields

$$
\begin{aligned}
CC_3 &= \left| \sum_{m} \exp(-i2\pi \underline{n}m/N)\, C(m-L)\, \chi(m, n-\underline{n}) \right|^2 \\
&= \left| \sum_{p} \exp(-i2\pi \underline{n}p/N)\, C(p)\, \chi(L+p, n-\underline{n}) \right|^2 .
\end{aligned}
\tag{38}
$$

Finally, as noted under equation (34) and using the evenness of real correlation $C(m)$, the covariance of interest can be modified into the two following equivalent forms:

$$
\begin{aligned}
\operatorname{cov}\{\mathbf{q}(n), \mathbf{p}(\underline{n})\} &= \left| \sum_{m} \exp(i2\pi \underline{n}m/N)\, C(m)\, \chi(L-m, n-\underline{n}) \right|^2 \\
&+ \left| \sum_{m} \exp(i2\pi \underline{n}m/N)\, C(m)\, \chi(L+m, n+\underline{n}) \right|^2 \\
&= \left| \sum_{m} \exp(-i2\pi \underline{n}m/N)\, C(m)\, \chi(L+m, n-\underline{n}) \right|^2 \\
&+ \left| \sum_{m} \exp(i2\pi \underline{n}m/N)\, C(m)\, \chi(L+m, n+\underline{n}) \right|^2 .
\end{aligned}
\tag{39}
$$

By setting $L = 0$, there follows

$$\mathrm{cov}\{\mathbf{q}(n),\mathbf{q}(\underline{n})\} = \left|\sum_m \exp(-i2\pi \underline{n} m/N)\, C(m)\, \chi(m,n-\underline{n})\right|^2 + \left|\sum_m \exp(i2\pi \underline{n} m/N)\, C(m)\, \chi(m,n+\underline{n})\right|^2 . \tag{40}$$

Finally,

$$\mathrm{var}\{\mathbf{q}(n)\} = \left|\sum_m \exp(-i2\pi\, n\, m/N)\, C(m)\, \chi(m,0)\right|^2 + \left|\sum_m \exp(i2\pi\, n\, m/N)\, C(m)\, \chi(m,2n)\right|^2 . \tag{41}$$

When the data correlation $C(m)$ is specialized to $\delta(m)$, namely, white Gaussian input noise, these results reduce to equations (18) through (20).

RELATION TO CORRELATION ESTIMATES OF THE INPUT DATA

Correlation estimates of the input data $\{\mathbf{g}(k)\}$ can be obtained from the magnitude-squared spectral estimates $\{\mathbf{q}(n)\}$ according to an inverse FFT:

$$\mathbf{r}(m) = \frac{1}{N}\sum_{n=0}^{N-1} \exp(i2\pi\, nm/N)\, \mathbf{q}(n) \ \text{ for } m = 0 : N-1. \tag{42}$$

Reference to equations (2) and (5) reveals that

$$\mathbf{q}(N-n) = \mathbf{q}(n) \ \text{ for } n = 1 : N-1. \tag{43}$$

Use of this relation in equation (42) shows that $\mathbf{r}(m)$ is real for all m and that

$$\mathbf{r}(N-m) = \mathbf{r}(m) \ \text{ for } m = 1 : N-1. \tag{44}$$

Then, it is only necessary to consider correlation estimates

$$\{\mathbf{r}(m)\} \ \text{ for } m = 0 : N/2. \tag{45}$$

Let $M = N/2 + 1$ and define vectors

$$\mathbf{q} = [\mathbf{q}(0)\ \ \mathbf{q}(1) \cdots \mathbf{q}(N-1)]^T,$$

$$e(m) = \frac{1}{N}[1\ \ \exp(i2\pi\, m/N) \cdots \exp\{i2\pi(N-1)m/N\}]^T \ \text{ for } m = 0 : N/2, \tag{46}$$

$$\mathbf{r} = [\mathbf{r}(0)\ \ \mathbf{r}(1) \cdots \mathbf{r}(N/2)]^T,$$

and $N \times M$ matrix

$$e = [e(0)\ \ e(1) \cdots e(N/2)]. \tag{47}$$

Then, by reference to equation (42),

$$\mathbf{r} = e^T\ \mathbf{q}. \tag{48}$$

The mean of correlation-estimate vector $\mathbf{r}$ is

$$E\{\mathbf{r}\} = e^T\ E\{\mathbf{q}\} \tag{49}$$

in terms of the mean of spectral vector $\mathbf{q}$, given by equations (28) or (32). Let the covariance matrix of $N \times 1$ vector $\mathbf{q}$ be denoted by C_{qq}; this quantity is available from equation (40). The covariance of vector $\mathbf{r}$ is then

$$\operatorname{cov}\{\mathbf{r}\} = E\{[\mathbf{r} - E\{\mathbf{r}\}]\,[\mathbf{r} - E\{\mathbf{r}\}]^T\} = e^T\ C_{qq}\ e \tag{50}$$

by use of equation (48). The matrix C_{qq} is $N \times N$ and is real, while matrix e is $N \times M$ and is complex; see equations (46) and (47). Nevertheless, the $M \times M$ covariance matrix of $\mathbf{r}$, obtained via equation (50), is purely real. Recall that $M = N/2 + 1$.

Another set of correlation estimates can be obtained from the alternative spectral estimates in equation (6), namely,

$$\mathbf{s}(m) = \frac{1}{N} \sum_{\underline{n}=0}^{N-1} \exp(i2\pi\, m\underline{n}/N)\ \mathbf{p}(\underline{n}) \ \text{ for } m = 0 : N/2. \tag{51}$$

The $M \times M$ covariance matrix between sets (45) and (51) is given by

$$\operatorname{cov}\{\mathbf{r},\mathbf{s}\} = e^T\ C_{qp}\ e, \tag{52}$$

where covariance matrix C_{qp} is available in equation (39), and is $N \times N$ and real.

A shortcut is possible by observing that equations (42) through (44) can be combined to yield

$$\mathbf{r}(m) = \frac{1}{N}[\mathbf{q}(0) + (-1)^m \, \mathbf{q}(N/2) + 2 \sum_{n=1}^{N/2-1} \cos(2\pi \, nm/N) \, \mathbf{q}(n)] \quad \text{for } m = 0 : N/2, \tag{53}$$

which involves only real quantities. Define $M \times 1$ vectors $(M = N/2 + 1)$

$$\begin{aligned} f(m) &= \frac{1}{N}[1 \;\; 2\cos(2\pi \, m/N) \cdots 2\cos\{2\pi \, m(N/2-1)/N\} \;\; (-1)^m]^T \quad \text{for } m = 0 : N/2, \\ \widetilde{\mathbf{q}} &= [\mathbf{q}(0) \;\; \mathbf{q}(1) \cdots \mathbf{q}(N/2)]^T, \quad \widetilde{\mathbf{p}} = [\mathbf{p}(0) \;\; \mathbf{p}(1) \cdots \mathbf{p}(N/2)]^T, \end{aligned} \tag{54}$$

and $M \times M$ matrix

$$f = [f(0) \;\; f(1) \cdots f(N/2)]. \tag{55}$$

Then,

$$\mathbf{r} = f^T \, \widetilde{\mathbf{q}} \text{ and } \mathbf{s} = f^T \, \widetilde{\mathbf{p}}. \tag{56}$$

The covariance matrix of vectors **r** and **s** then becomes

$$\mathrm{cov}\{\mathbf{r}, \mathbf{s}\} = f^T \, C_{\widetilde{q}\widetilde{p}} \, f. \tag{57}$$

Covariance matrix $C_{\widetilde{q}\widetilde{p}}$ is $M \times M$ and real, while matrix f is also $M \times M$ and real; here, $M = N/2 + 1$. Thus, equation (57) has two advantages relative to equation (52), namely, smaller matrices (by a factor of two) and purely real quantities. The information required for construction of the covariance matrix $C_{\widetilde{q}\widetilde{p}}$ in equation (57) has been presented in equation (39).

More explicit relations for these correlation-estimate statistics are presented in sections 3, 4, and 5. Also, the statistics for the *complex* spectral estimates are presented in section 6.

3. STATISTICS OF CORRELATION ESTIMATES FOR WHITE DATA AND $\mathrm{N} \geq \mathrm{K}$

Sequence $\{\mathbf{g}(k)\}$ consists of independent, real, Gaussian, random variables (RVs) with zero mean and unit variance. Real weighting sequence $\{w(k)\}$ is defined for all integer k, but is nonzero only for $k = 1:K$. A complex spectral estimate in frequency bin n is obtained according to

$$\mathbf{z}(n) = \sum_k \mathbf{g}(k)\, w(k) \exp(-i2\pi\, n\, k/N) \quad \text{for } n = 0:N-1. \tag{58}$$

The infinite sum on k is automatically terminated to K terms by the weighting; this isolates a segment of the data stream $\{\mathbf{g}(k)\}$. The integer N governs the frequency spacing of the complex spectral estimates $\{\mathbf{z}(n)\}$. Integer N is presumed even.

In addition, another set of complex spectral estimates is obtained from a lagged data set according to

$$\mathbf{y}(n) = \sum_k \mathbf{g}(k)\, w(k-L) \exp[-i2\pi\, n\,(k-L)/N] = \sum_k \mathbf{g}(k+L)\, w(k) \exp(-\,i2\pi\, nk/N), \tag{59}$$

where L is the amount of lag between the two sets of data and is arbitrary. If lag L is greater than or equal to weighting duration K, the two sets of spectral estimates in equations (58) and (59) are statistically independent of each other, because they involve different sets of independent Gaussian RVs. Interest here will be concentrated on the case $0 \leq L < K$, meaning that the two segmented data sets encountered in equations (58) and (59) are statistically dependent on each other, thereby making complex RVs $\{\mathbf{z}(n)\}$ and $\{\mathbf{y}(n)\}$ statistically dependent on each other. It is also assumed that $N \geq K$; thus, the various integers of the following analysis satisfy

$$0 \leq L < K \leq N, \tag{60}$$

but L can be arbitrary in general.

The correlation estimates of interest are defined according to inverse FFTs as

$$\mathbf{r}(m) = \frac{1}{N} \sum_{n=0}^{N-1} |\mathbf{z}(n)|^2 \exp(i2\pi\, nm/N) \quad \text{for } m = 0:N-1 \tag{61}$$

and

$$\mathbf{q}(m) = \frac{1}{N}\sum_{n=0}^{N-1} |\mathbf{y}(n)|^2 \exp(i2\pi\, nm/N) \quad \text{for} \quad m = 0 : N-1. \tag{62}$$

By using the property from equation (58) that

$$\mathbf{z}(N-n) = \mathbf{z}^*(n) \quad \text{for} \quad n = 1 : N-1, \tag{63}$$

it can be shown that all the correlation estimates $\{\mathbf{r}(m)\}$ in equation (61) are real. It then also follows that

$$\mathbf{r}(N-m) = \mathbf{r}(m) \quad \text{for} \quad m = 1 : N-1. \tag{64}$$

Therefore, it is only necessary to consider the correlation estimates

$$\{\mathbf{r}(m)\} \quad \text{for} \quad m = 0 : N/2 \tag{65}$$

in the following. A similar set of properties holds for the alternative set of correlation estimates $\{\mathbf{q}(m)\}$ in equation (62).

The autocorrelation of the white data sequence $\{\mathbf{g}(k)\}$ in equation (58) is

$$E\{\mathbf{g}(k)\,\mathbf{g}(j)\} = \delta(k-j), \tag{66}$$

where the Kronecker delta function is defined here as

$$\delta(k) = \begin{Bmatrix} 1 \text{ for } k = 0 \\ 0 \text{ otherwise} \end{Bmatrix}. \tag{67}$$

Then, from equation (58), the mean of magnitude-squared spectral estimate $\mathbf{z}(n)$ is immediately

$$E\{|\mathbf{z}(n)|^2\} = \sum_k w^2(k) \quad \text{for } n = 0 : N-1. \tag{68}$$

Therefore, the mean of correlation estimate $\mathbf{r}(m)$ in equation (61) is

$$E\{\mathbf{r}(m)\} = \sum_k w^2(k)\; \frac{1}{N}\sum_{n=0}^{N-1} \exp(i2\pi\, nm/N) = \sum_k w^2(k)\;\; \delta_N(m), \tag{69}$$

where

$$\delta_N(m) \equiv \begin{Bmatrix} 1 \;\text{ for }\; m = 0, \pm N, \pm 2N, \ldots \\ 0 \;\text{ otherwise} \end{Bmatrix} \tag{70}$$

is the periodic Kronecker delta function. Therefore, use of equations (65) and (70) reveals that equation (69) can be simplified to

$$E\{\mathbf{r}(m)\} = \left\{ \begin{array}{ll} \sum_{k} w^2(k) & \text{for } m = 0 \\ 0 & \text{for } m = 1 : N/2 \end{array} \right\}. \tag{71}$$

Of course, correlation estimates $\{\mathbf{q}(m)\}$ in equation (62) have the same mean values.

The next statistical quantity of interest is the covariance between correlation estimates $\mathbf{r}(m)$ and $\mathbf{q}(\underline{m})$ in equations (61) and (62), where integers m and $\underline{m}$ may be equal or not. The crosscorrelation of these two RVs is

$$CC \equiv E\{\mathbf{r}(m)\, \mathbf{q}(\underline{m})\} = \frac{1}{N^2} \sum_{n,\underline{n}=0}^{N-1} E\{|\mathbf{z}(n)|^2\, |\mathbf{y}(\underline{n})|^2\} \exp[i2\pi(n\,m + \underline{n}\,\underline{m})/N]. \tag{72}$$

Denote the ensemble average in equation (72) as A. Then, from equations (58) and (59),

$$A = \sum_{kj\underline{k}\underline{j}} E\{\mathbf{g}(k)\, \mathbf{g}(j)\, \mathbf{g}(\underline{k} + L)\, \mathbf{g}(\underline{j} + L)\}\, w(k)\, w(j)\, w(\underline{k})\, w(\underline{j}) \\ \times \exp[-i2\pi\, \{n(k - j) + \underline{n}(\underline{k} - \underline{j})\}/N]. \tag{73}$$

Denote the ensemble average in equation (73) as B. Then, using the Gaussian character of data sequence $\{\mathbf{g}(k)\}$ and equation (66), this ensemble average can be expressed as the sum of three terms, namely,

$$B = \delta(k - j)\, \delta(\underline{k} - \underline{j}) + \delta(k - \underline{k} - L)\, \delta(j - \underline{j} - L) + \delta(k - \underline{j} - L)\, \delta(j - \underline{k} - L). \tag{74}$$

Substitution of the first product of Kronecker delta functions into equation (73) yields the first component of the average A as

$$A_1 = \sum_{k,\underline{k}} w^2(k)\, w^2(\underline{k}) = \left(\sum_{k} w^2(k) \right)^2. \tag{75}$$

Use of this result in equation (72) yields the first component of the crosscorrelation CC as

$$CC_1 = \left(\sum_{k} w^2(k) \right)^2 \frac{1}{N^2} \sum_{n,\underline{n}=0}^{N-1} \exp[i2\pi(n\,m + \underline{n}\,\underline{m})/N] = \left(\sum_{k} w^2(k) \right)^2 \delta_N(m)\, \delta_N(\underline{m}) \\ = E\{\mathbf{r}(m)\}\, E\{\mathbf{q}(\underline{m})\}, \tag{76}$$

by reference to equation (69) and the comment under equation (71). Thus, the sum of the two remaining components of CC in equation (72) will be the covariance of RVs $\mathbf{r}(m)$ and $\mathbf{q}(\underline{m})$.

Substitution of the second product of equation (74) into equation (73) yields the second component of average A as

$$A_2 = \sum_{k,j} w(k)\, w(k-L)\, w(j)\, w(j-L) \exp[-i2\pi(n+\underline{n})(k-j)/N]. \tag{77}$$

Use of this result in equation (72) yields the second component of the crosscorrelation CC as

$$\begin{aligned} CC_2 &= \sum_{k,j} w(k)\, w(k-L)\, w(j)\, w(j-L) \\ &\quad \times \frac{1}{N^2} \sum_{n,\underline{n}=0}^{N-1} \exp[-i2\pi(n+\underline{n})(k-j)/N + i2\pi(n\,m+\underline{n}\,\underline{m})/N] \\ &= \sum_{k,j} w(k)\, w(k-L)\, w(j)\, w(j-L)\, \delta_N(m-k+j)\, \delta_N(\underline{m}-k+j). \end{aligned} \tag{78}$$

The second periodic Kronecker delta function is nonzero when

$$j = k - \underline{m} + p\,N, \tag{79}$$

where p is an arbitrary integer, positive or negative or zero. Then, there follows

$$CC_2 = \sum_{k} w(k)\, w(k-L) \sum_{p} w(k-\underline{m}+p\,N)\, w(k-\underline{m}+p\,N-L)\, \delta_N(m-\underline{m}+p\,N). \tag{80}$$

The $p\,N$ term in the periodic δ_N function can obviously be dropped. Since both m and $\underline{m}$ are limited according to equation (65), the subscript N can also be dropped.

Now, define the function

$$h(k,L) = w(k)\, w(k-L), \tag{81}$$

which is of length $K-L$. Then, equation (80) yields

$$\begin{aligned} CC_2 &= \delta(m-\underline{m}) \sum_{k} h(k,L) \sum_{p} h(k-m+p\,N, L) \\ &= \delta(m-\underline{m}) \sum_{k} h(k,L)\,[h(k-m,L) + h(k-m+N,L) + h(k-m-N,L) + \cdots]. \end{aligned} \tag{82}$$

Also, define the function

$$H(m,L) = \sum_k h(k,L)\, h(k-m,L) \text{ for all } L, m. \tag{83}$$

Function $H(m,L)$ is even in m and

$$H(m,L) \neq 0 \text{ only for } |m| < K - L. \tag{84}$$

Equation (82) can now be expressed as

$$CC_2 = \delta(m - \underline{m})\,[H(m,L) + H(m-N,L) + H(m+N,L) + \cdots]. \tag{85}$$

But since $N \geq K$ and $0 \leq m \leq N/2$, the only terms that need to be kept are

$$CC_2 = \delta(m - \underline{m})\,[H(m,L) + H(m-N,L)] \text{ for } m = 0 : N/2, \tag{86}$$

where some of the terms may be zero.

Substitution of the third product of equation (74) into equation (73) yields the third component of average A as

$$A_3 = \sum_{k,j} w(k)\, w(k-L)\, w(j)\, w(j-L) \exp[-i2\pi\,(n - \underline{n})(k-j)/N]. \tag{87}$$

The only difference with equation (77) is in the polarity of the $\underline{n}$ term. Use of this result in equation (72) yields, after simplifications similar to equation (78), the third component of the crosscorrelation CC as

$$CC_3 = \sum_{k,j} h(k,L)\, h(j,L)\, \delta_N(m-k+j)\, \delta_N(\underline{m}+k-j). \tag{88}$$

Definition (81) has also been employed here. At this point, the argument is identical to that used above in equations (79) through (86), with the end result that

$$CC_3 = \delta(m)\,\delta(\underline{m})\,H(0,L) + \delta(m-N/2)\,\delta(\underline{m}-N/2)\,2\,H(N/2,L). \tag{89}$$

When results (86) and (89) are added together, the end result is the desired covariance, as noted under equation (76); namely,

$$\begin{aligned} \text{cov}\{\mathbf{r}(m), \mathbf{q}(\underline{m})\} = {} & \delta(m-\underline{m})\,[H(m,L) + H(m-N,L)] \\ & + \delta(m)\,\delta(\underline{m})\,H(0,L) + \delta(m-N/2)\,\delta(\underline{m}-N/2)\,2\,H(N/2,L) \end{aligned} \tag{90}$$

for $m, \underline{m} = 0 : N/2$. In more specific terms,

$$
\begin{aligned}
&\operatorname{cov}\{\mathbf{r}(0),\mathbf{q}(0)\} = 2\,H(0,L),\\
&\operatorname{cov}\{\mathbf{r}(N/2),\mathbf{q}(N/2)\} = 4\,H(N/2,L),\\
&\operatorname{cov}\{\mathbf{r}(m),\mathbf{q}(m)\} = H(m,L) + H(m-N,L) \ \text{ for } m = 1 : N/2-1,\\
&\operatorname{cov}\{\mathbf{r}(m),\mathbf{q}(\underline{m})\} = 0 \ \text{ for } m \neq \underline{m}.
\end{aligned}
\tag{91}
$$

The function $H(m,L)$ is defined in equation (83) as the autocorrelation of sequence $\{h(k,L)\}$. This latter sequence is defined in equation (81) as the product of the original weighting $\{w(k)\}$ and a delayed version by L units. An example MATLAB program for equation (91) is given in figure 2.

```
function stat_pmb
clear all
close all hidden

global K L xh
K=1024;                  % Length of weighting
L=512;                   % Lag (shift) of segment; 0 <= L < K
N=2048;                  % FFT size; K <= N
% w=hann(K);
% w=hann(K+1); w=w(2:K+1);
w=hann(K+2); w=w(2:K+1); % Eliminate zero weights

average=sum(w.^2);

h=w(L+1:K).*w(1:K-L);
xh=xcorr(h);
cov00=2*H(0);                 % cov{r(0),q(0)}
cov2=zeros(N/2,1);
for m=1:N/2-1
     cov2(m)=H(m)+H(m-N); % cov{r(m),q(m)} for m=1:N/2
end
cov2(N/2)=4*H(N/2);

keyboard

function w = H(m)
global K L xh
M=K-L;
if(abs(m)>=M)
     w=0;
else
     w=xh(M+m);
end
```

Figure 2. MATLAB Example for Equation (91)

4. STATISTICS OF CORRELATION ESTIMATES FOR NON-WHITE DATA AND N ≥ 2K

The results in the previous section were obtained by means of frequency domain relations. It is now more convenient, for non-white data, to work solely in the time domain. From equations (58), (61), and (70), there follows the correlation estimate

$$\begin{aligned}\mathbf{r}(m) &= \frac{1}{N}\sum_{n=0}^{N-1} \exp(i2\pi\, mn/N)\left|\sum_k \mathbf{g}(k)\; w(k)\; \exp(-i2\pi\, nk/N)\right|^2 \\ &= \sum_{k,\underline{k}} \mathbf{g}(k)\,\mathbf{g}(\underline{k})\; w(k)\; w(\underline{k})\,\frac{1}{N}\sum_{n=0}^{N-1} \exp[i2\pi(m-k+\underline{k})\,n/N] \\ &= \sum_{k,\underline{k}} \mathbf{g}(k)\,\mathbf{g}(\underline{k})\; w(k)\; w(\underline{k})\;\delta_N(m-k+\underline{k}).\end{aligned} \tag{92}$$

For a given value of k, the only values of $\underline{k}$ that can contribute are

$$\underline{k} = k - m \text{ and } \underline{k} = k - m + N, \tag{93}$$

presuming that $N \geq K$ and that m is limited to the range $0 : N/2$, as in equation (65). Equation (92) then simplifies to

$$\mathbf{r}(m) = \sum_k \mathbf{g}(k)\,\mathbf{g}(k-m)\; w(k)\; w(k-m) + \sum_k \mathbf{g}(k)\,\mathbf{g}(k-m+N)\; w(k)\; w(k-m+N). \tag{94}$$

When the product

$$\mathbf{p}(k) = \mathbf{g}(k)\; w(k) \tag{95}$$

is defined, equation (94) can be written as

$$\mathbf{r}(m) = \sum_k \mathbf{p}(k)\,\mathbf{p}(k-m) + \sum_k \mathbf{p}(k)\,\mathbf{p}(k-m+N), \tag{96}$$

which is the sum of two autocorrelations of $\mathbf{p}(k)$, one of which is shifted by N units. Thus, wraparound can be present in correlation estimate $\mathbf{r}(m)$ when the condition $N \geq K$ is imposed.

In this section, interest is directed at the more stringent condition

$$N \geq 2K \text{ and } m = 0 : N/2, \tag{97}$$

for which the last summation in equation (96) is zero because the two **p** functions involved cannot overlap; that is, there is *no* wraparound in the time-delay domain. Thus, the case of immediate interest is given by the two random variables

$$\begin{aligned} \mathbf{r}(m) &= \sum_k \mathbf{g}(k)\, \mathbf{g}(k-m)\, w(k)\, w(k-m), \\ \mathbf{q}(\underline{m}) &= \sum_{\underline{k}} \mathbf{g}(\underline{k}+L)\, \mathbf{g}(\underline{k}+L-\underline{m})\, w(\underline{k})\, w(\underline{k}-\underline{m}), \end{aligned} \tag{98}$$

where both m and $\underline{m}$ are limited to the range $0 : N/2$.

For non-white input data $\{\mathbf{g}(k)\}$, and by use of equations (27) and (30), the means of correlation estimates $\mathbf{r}(m)$ and $\mathbf{q}(\underline{m})$ are immediately available as

$$\begin{aligned} E\{\mathbf{r}(m)\} &= \sum_k C(m)\, w(k)\, w(k-m) = C(m)\, \phi_w(m), \\ E\{\mathbf{q}(\underline{m})\} &= \sum_{\underline{k}} C(\underline{m})\, w(\underline{k})\, w(\underline{k}-\underline{m}) = C(\underline{m})\, \phi_w(\underline{m}). \end{aligned} \tag{99}$$

If the length K of weighting $\{w(k)\}$ is large, correlation estimates (98) will be only slightly biased near their origins.

The crosscorrelation of the two correlation estimates in equation (98) is

$$\begin{aligned} CC \equiv E\{\mathbf{r}(m)\, \mathbf{q}(\underline{m})\} = \sum_{k,\underline{k}} & E\{\mathbf{g}(k)\, \mathbf{g}(k-m)\, \mathbf{g}(\underline{k}+L)\, \mathbf{g}(\underline{k}+L-\underline{m})\} \\ & \times w(k)\, w(k-m)\, w(\underline{k})\, w(\underline{k}-\underline{m}). \end{aligned} \tag{100}$$

Using the Gaussian character of the data $\{\mathbf{g}(k)\}$, the ensemble average is given by

$$C(m)\, C(\underline{m}) + C(k-\underline{k}-L)\, C(k-\underline{k}-L-m+\underline{m}) + C(k-\underline{k}-L+\underline{m})\, C(k-\underline{k}-L-m). \tag{101}$$

Substitution of the first product of equation (101) into equation (100) yields the first component of crosscorrelation CC as

$$CC_1 = C(m)\, C(\underline{m}) \sum_{k,\underline{k}} w(k)\, w(k-m)\, w(\underline{k})\, w(\underline{k}-\underline{m})$$
$$= C(m)\, C(\underline{m})\, \phi_w(\mathrm{m})\, \phi_w(\underline{\mathrm{m}}) = E\{\mathbf{r}(m)\}\, E\{\mathbf{q}(\underline{m})\}, \tag{102}$$

by reference to equation (99). Therefore, the sum of the two remaining terms of CC in equation (100) will yield the covariance of correlation estimates $\mathbf{r}(m)$ and $\mathbf{q}(\underline{m})$.

The second component of CC is obtained by substituting the second product of equation (101) into equation (100):

$$CC_2 = \sum_{k,\underline{k}} C(k-\underline{k}-L)\, C(k-\underline{k}-L-m+\underline{m})\, w(k)\, w(k-m)\, w(\underline{k})\, w(\underline{k}-\underline{m})$$
$$= \sum_{p} C(p-L)\, C(p-L-m+\underline{m}) \sum_{k} w(k)\, w(k-m)\, w(k-p)\, w(k-p-\underline{m}) \tag{103}$$
$$= \sum_{p} C(p-L)\, C(p-L-m+\underline{m})\, \theta(m,p,\underline{m}+p),$$

where substitution $p = k - \underline{k}$ was made, and the third-order autocorrelation of the weighting $\{w(k)\}$ is defined as

$$\theta(m,n,p) = \sum_{k} w(k)\, w(k-m)\, w(k-n)\, w(k-p). \tag{104}$$

The approximate evaluation of this third-order correlation is presented in appendix B.

The third component of CC is obtained by substituting the third product of equation (101) into equation (100):

$$CC_3 = \sum_{k,\underline{k}} C(k-\underline{k}-L+\underline{m})\, C(k-\underline{k}-L-m)\, w(k)\, w(k-m)\, w(\underline{k})\, w(\underline{k}-\underline{m})$$
$$= \sum_{p} C(p-L+\underline{m})\, C(p-L-m) \sum_{k} w(k)\, w(k-m)\, w(k-p)\, w(k-p-\underline{m}) \tag{105}$$
$$= \sum_{p} C(p-L+\underline{m})\, C(p-L-m)\, \theta(m,p,\underline{m}+p).$$

Therefore, according to the observation under equation (102), the final covariance of interest is

$$\mathrm{cov}\{\mathbf{r}(m),\mathbf{q}(\underline{m})\} = \sum_{k} [C(k-L)\, C(k-L-m+\underline{m}) + C(k-L+\underline{m})\, C(k-L-m)]\, \theta(m,k,\underline{m}+k). \tag{106}$$

As special cases,

$$\operatorname{cov}\{\mathbf{r}(m),\mathbf{r}(\underline{m})\} = \sum_k [C(k)\, C(k-m+\underline{m}) + C(k+\underline{m})\, C(k-m)]\, \theta(m,k,\underline{m}+k) \tag{107}$$

and

$$\operatorname{var}\{\mathbf{r}(m)\} = \sum_k [C^2(k) + C(k+m)\, C(k-m)]\, \theta(m,k,m+k). \tag{108}$$

5. STATISTICS OF CORRELATION ESTIMATES FOR NON-WHITE DATA AND N ≥ K

In this section, the less stringent condition, $N \geq K$, is considered; the pertinent relation for the correlation estimate $\mathbf{r}(m)$ is then equation (94):

$$\begin{aligned}\mathbf{r}(m) &= \sum_{k} \mathbf{g}(k)\,\mathbf{g}(k-m)\,w(k)\,w(k-m) + \sum_{k} \mathbf{g}(k)\,\mathbf{g}(k-m+N)\,w(k)\,w(k-m+N) \\ &\equiv \mathbf{r}_1(m) + \mathbf{r}_2(m),\end{aligned} \tag{109}$$

which exhibits time-delay-domain wraparound in the second term. Also, the additional correlation estimate $\mathbf{q}(\underline{m})$ now takes the form

$$\begin{aligned}\mathbf{q}(\underline{m}) &= \sum_{\underline{k}} \mathbf{g}(\underline{k}+L)\,\mathbf{g}(\underline{k}+L-\underline{m})\,w(\underline{k})\,w(\underline{k}-\underline{m}) \\ &+ \sum_{\underline{k}} \mathbf{g}(\underline{k}+L)\,\mathbf{g}(\underline{k}+L-\underline{m}+N)\,w(\underline{k})\,w(\underline{k}-\underline{m}+N) \\ &\equiv \mathbf{q}_1(\underline{m}) + \mathbf{q}_2(\underline{m}),\end{aligned} \tag{110}$$

which also exhibits wraparound in the second term.

The means of these two correlation estimates are immediately available as

$$\begin{aligned}E\{\mathbf{r}(m)\} &= C(m)\,\phi_w(m) + C(m-N)\,\phi_w(m-N), \\ E\{\mathbf{q}(\underline{m})\} &= C(\underline{m})\,\phi_w(\underline{m}) + C(\underline{m}-N)\,\phi_w(\underline{m}-N),\end{aligned} \tag{111}$$

upon use of equations (27) and (30).

The crosscorrelation of estimates $\mathbf{r}(m)$ and $\mathbf{q}(\underline{m})$ is given by

$$\begin{aligned}CC = E\{\mathbf{r}(m)\,\mathbf{q}(\underline{m})\} &= E\{\mathbf{r}_1(m)\,\mathbf{q}_1(\underline{m})\} + E\{\mathbf{r}_1(m)\,\mathbf{q}_2(\underline{m})\} \\ &+ E\{\mathbf{r}_2(m)\,\mathbf{q}_1(\underline{m})\} + E\{\mathbf{r}_2(m)\,\mathbf{q}_2(\underline{m})\} \equiv P + Q + R + S.\end{aligned} \tag{112}$$

The first term, P, is available from equations (109) and (110) as

$$P = \sum_{k,\underline{k}} E\{\mathbf{g}(k)\,\mathbf{g}(k-m)\,\mathbf{g}(\underline{k}+L)\,\mathbf{g}(\underline{k}+L-\underline{m})\}\,w(k)\,w(k-m)\,w(\underline{k})\,w(\underline{k}-\underline{m}). \tag{113}$$

Using the Gaussian character of data $\{\mathbf{g}(k)\}$, the ensemble average is given by

$$C(m)\,C(\underline{m}) + C(k-\underline{k}-L)\,C(k-\underline{k}-L-m+\underline{m}) + C(k-\underline{k}-L+\underline{m})\,C(k-\underline{k}-L-m). \tag{114}$$

Substitution of the first product of equation (114) into equation (113) yields the first term of P as

$$P_1 = C(m)\, C(\underline{m})\, \phi_w(m)\, \phi_w(\underline{m}), \tag{115}$$

by a now familiar procedure. Substitution of the second and third products of equation (114) into equation (113), and use of the change of variables $p = k - \underline{k}$, leads to the respective terms

$$\begin{aligned} P_2 &= \sum_p C(p-L)\, C(p-L-m+\underline{m})\, \theta(m,p,\underline{m}+p), \\ P_3 &= \sum_p C(p-L+\underline{m})\, C(p-L-m)\, \theta(m,p,\underline{m}+p). \end{aligned} \tag{116}$$

For the Q average in equation (112), the ensemble average and the three components of Q are

$$\begin{aligned} &C(m)\, C(\underline{m}-N) + C(k-\underline{k}-L)\, C(k-\underline{k}-L-m+\underline{m}-N) \\ &\qquad + C(k-\underline{k}-L+\underline{m}-N)\, C(k-\underline{k}-L-m), \end{aligned} \tag{117}$$

and

$$\begin{aligned} Q_1 &= C(m)\, C(\underline{m}-N)\, \phi_w(m)\, \phi_w(\underline{m}-N), \\ Q_2 &= \sum_p C(p-L)\, C(p-L-m+\underline{m}-N)\, \theta(m,p,\underline{m}+p-N), \\ Q_3 &= \sum_p C(p-L+\underline{m}-N)\, C(p-L-m)\, \theta(m,p,\underline{m}+p-N). \end{aligned} \tag{118}$$

The corresponding results for the R average in equation (112) are

$$\begin{aligned} &C(m-N)\, C(\underline{m}) + C(k-\underline{k}-L)\, C(k-\underline{k}-L-m+\underline{m}+N) \\ &\qquad + C(k-\underline{k}-L+\underline{m})\, C)k-\underline{k}-L-m+N), \end{aligned} \tag{119}$$

and

$$\begin{aligned} R_1 &= C(m-N)\, C(\underline{m})\, \phi_w(m-N)\, \phi_w(\underline{m}), \\ R_2 &= \sum_p C(p-L)\, C(p-L-m+\underline{m}+N)\, \theta(m-N,p,\underline{m}+p), \\ R_3 &= \sum_p C(p-L+\underline{m})\, C(p-L-m+N)\, \theta(m-N,p,\underline{m}+p). \end{aligned} \tag{120}$$

Finally, the corresponding results for the S average in equation (112) are

$$\begin{aligned} &C(m-N)\, C(\underline{m}-N) + C(k-\underline{k}-L)\, C(k-\underline{k}-L-m+\underline{m}-N) \\ &\qquad + C(k-\underline{k}-L+\underline{m}-N)\, C(k-\underline{k}-L-m+N), \end{aligned} \tag{121}$$

and

$$S_1 = C(m-N)\,C(\underline{m}-N)\,\phi_w(m-N)\,\phi_w(\underline{m}-N),$$
$$S_2 = \sum_p C(p-L)\,C(p-L-m+\underline{m})\,\theta(m-N,p,\underline{m}+p-N), \tag{122}$$
$$S_3 = \sum_p C(p-L+\underline{m}-N)\,C(p-L-m+N)\,\theta(m-N,p,\underline{m}+p-N).$$

The sum of the first terms of these quantities, namely, $P_1+Q_1+R_1+S_1$, is recognized as the product of the means in equation (111). Therefore, the sum of all the remaining terms in equations (116), (118), (120), and (122) yields the desired covariance:

$$\begin{aligned}
&\operatorname{cov}\{\mathbf{r}(m),\mathbf{q}(\underline{m})\}\\
&=\sum_p [C(p-L)\,C(p-L-m+\underline{m})+C(p-L+\underline{m})\,C(p-L-m)]\,\theta(m,p,\underline{m}+p)\\
&+\sum_p [C(p-L)\,C(p-L-m+\underline{m}-N)+C(p-L+\underline{m}-N)\,C(p-L-m)]\,\theta(m,p,\underline{m}+p-N)\\
&+\sum_p [C(p-L)\,C(p-L-m+\underline{m}+N)+C(p-L+\underline{m})\,C(p-L-m+N)]\,\theta(m-N,p,\underline{m}+p)\\
&+\sum_p [C(p-L)\,C(p-L-m+\underline{m})+C(p-L+\underline{m}-N)\,C(p-L-m+N)]\,\theta(m-N,p,\underline{m}+p-N).
\end{aligned} \tag{123}$$

At this point, it is convenient to define the function

$$\begin{aligned}
T(L,m,\underline{m}) \equiv \sum_p &[C(p-L)\,C(p-L-m+\underline{m})\\
&+C(p-L+\underline{m})\,C(p-L-m)]\,\theta(m,p,\underline{m}+p).
\end{aligned} \tag{124}$$

Then, the covariance in equation (123) can be expressed in a compact symmetric form as

$$\operatorname{cov}\{\mathbf{r}(m),\mathbf{q}(\underline{m})\} = T(L,m,\underline{m})+T(L,m,\underline{m}-N)+T(L,m-N,\underline{m})+T(L,m-N,\underline{m}-N). \tag{125}$$

An alternative (slightly simpler) expression for function $T(L,m,\underline{m})$ is

$$T(L,m,\underline{m}) = \sum_k [C(k)\,C(k-m+\underline{m})+C(k+\underline{m})\,C(k-m)]\,\theta(m,k+L,\underline{m}+k+L). \tag{126}$$

As special cases of the above, there follows

$$\operatorname{cov}\{\mathbf{r}(m),\mathbf{r}(\underline{m})\} = T(0,m,\underline{m})+T(0,m,\underline{m}-N)+T(0,m-N,\underline{m})+T(0,m-N,\underline{m}-N) \tag{127}$$

and

$$\operatorname{var}\{\mathbf{r}(m)\} = T(0,m,m)+T(0,m,m-N)+T(0,m-N,m)+T(0,m-N,m-N). \tag{128}$$

The function $T(L,m,\underline{m})$ in equation (124) involves

$$\begin{aligned}\theta(m,p,\underline{m}+p) &= \sum_k w(k)\, w(k-m)\, w(k-p)\, w(k-p-\underline{m}) \\ &= \sum_k h(k,m)\, h(k-p,\underline{m}) \equiv \underline{H}(p,m,\underline{m}),\end{aligned} \tag{129}$$

where auxiliary function

$$h(k,m) \equiv w(k)\, w(k-m). \tag{130}$$

The sequence $\{h(k,m)\}$ is a product of two weighting sequences, one of which is delayed by m units. Thus, $\theta(m,p,\underline{m}+p)$ can be interpreted as a crosscorrelation of $\{h(k,m)\}$ with $\{h(k,\underline{m})\}$, versus delay p. Then, the sum on p in equation (124) yields

$$\begin{aligned}T(L,m,\underline{m}) &= \sum_p \left[C(p-L)\, C(p-L-m+\underline{m}) + C(p-L+\underline{m})\, C(p-L-m)\right] \underline{H}(p,m,\underline{m}) \\ &= \sum_k \left[C(k)\, C(k-m+\underline{m}) + C(k+\underline{m})\, C(k-m)\right] \underline{H}(k+L,m,\underline{m}).\end{aligned} \tag{131}$$

For white input data $\{\mathbf{g}(k)\}$, $C(k)=\delta(k)$, yielding

$$T(L,m,\underline{m}) = \delta(m-\underline{m})\, \underline{H}(L,m,m) + \delta(m+\underline{m})\, \underline{H}(m+L,m,-m). \tag{132}$$

As a further special case for $m=\underline{m}=0$, there follows

$$\underline{H}(L,0,0) = \sum_k h(k,0)\, h(k-L,0) = \sum_k w^2(k)\, w^2(k-L). \tag{133}$$

More generally,

$$\underline{H}(L,m,m) = \sum_k h(k,m)\, h(k-L,m) \equiv H(L,m), \tag{134}$$

while

$$T(L,m,\underline{m}) = \delta(m-\underline{m})\, H(L,m), \text{ except for } m=\underline{m}=0. \tag{135}$$

6. STATISTICS OF COMPLEX SPECTRAL ESTIMATES

Section 2 treated the statistics of the magnitude-squared spectral estimates. Here, the quantities of interest are the *complex* spectral estimates

$$\mathbf{z}(n) = \sum_k \mathbf{g}(k)\, w(k) \exp(-i2\pi\, n\, k/N) \quad \text{for} \quad n = 0 : N-1 \tag{136}$$

and the lagged complex spectral estimates

$$\tilde{\mathbf{z}}(n) = \sum_k \mathbf{g}(k+L)\, w(k) \exp(-i2\pi\, n\, k/N) \quad \text{for} \quad n = 0 : N-1 \tag{137}$$

for white, Gaussian input $\{\mathbf{g}(k)\}$.

For later use, define window functions

$$W_1(n) = \sum_k w(k)\, w(k-L) \exp(-i\, 2\,\pi\, n\, k/N) \quad \text{for all } n \tag{138}$$

and

$$W_2(n) = \sum_k w^2(k) \exp(-i\, 2\,\pi\, n\, k/N) \quad \text{for all } n. \tag{139}$$

Both windows have period N in n while their real parts are even in n and their imaginary parts are odd in n.

The mean of $\mathbf{z}(n)$ is zero, while the two second-order moments of $\mathbf{z}(n)$ are

$$\begin{aligned} E\{\mathbf{z}^2(n)\} &= \sum_k w^2(k) \exp(-i4\,\pi\, n\, k/N) = W_2(2\,n), \\ E\{|\mathbf{z}(n)|^2\} &= \sum_k w^2(k) = W_2(0). \end{aligned} \tag{140}$$

The two second-order moments for separated frequency bins n and $\underline{n}$ are

$$\begin{aligned} E\{\mathbf{z}(n)\, \mathbf{z}(\underline{n})\} &= \sum_k w^2(k) \exp[-i\, 2\,\pi\,(n+\underline{n})\, k/N] = W_2(n+\underline{n}), \\ E\{\mathbf{z}(n)\, \mathbf{z}^*(\underline{n})\} &= \sum_k w^2(k) \exp[-i\, 2\,\pi\,(n-\underline{n})\, k/N] = W_2(n-\underline{n}). \end{aligned} \tag{141}$$

These equations give the complete covariance information about the complex spectral estimates $\{\mathbf{z}(n)\}$.

Let the complex spectral estimate be represented in terms of its real and imaginary parts according to $\mathbf{z}(n) = \mathbf{x}(n) + i\,\mathbf{y}(n)$. Then, $E\{\mathbf{x}(n)\} = 0,\ E\{\mathbf{y}(n)\} = 0$. Upon substitution into equations (141), and balancing of real and imaginary parts, there follows

$$\begin{aligned}
E\{\mathbf{x}(n)\,\mathbf{x}(\underline{n})\} &= \frac{1}{2} W_{2r}(n-\underline{n}) + \frac{1}{2} W_{2r}(n+\underline{n}), \\
E\{\mathbf{y}(n)\,\mathbf{y}(\underline{n})\} &= \frac{1}{2} W_{2r}(n-\underline{n}) - \frac{1}{2} W_{2r}(n+\underline{n}), \\
E\{\mathbf{x}(n)\,\mathbf{y}(\underline{n})\} &= -\frac{1}{2} W_{2i}(n-\underline{n}) + \frac{1}{2} W_{2i}(n+\underline{n}), \\
E\{\mathbf{x}(\underline{n})\,\mathbf{y}(n)\} &= \frac{1}{2} W_{2i}(n-\underline{n}) + \frac{1}{2} W_{2i}(n+\underline{n}).
\end{aligned} \tag{142}$$

These equations contain complete covariance information about the real and imaginary parts of the complex spectral estimates $\{\mathbf{z}(n)\}$. If $\underline{n}$ and n are switched in the fourth expression, the result is equal to the third expression, because W_{2i} is an odd function. If sequence $\{W_2(n)\}$ in equation (139) is computed by means of an N-point FFT, then replace $n \pm \underline{n}$ by $\mathrm{mod}(n \pm \underline{n}, N)$.

The second-order statistics of lagged estimate $\tilde{\mathbf{z}}(n)$ in equation (137) are identical to the corresponding terms presented above. Now, let n and $\underline{n}$ be different (or equal) integers. Then,

$$\begin{aligned}
E\{\mathbf{z}(n)\,\tilde{\mathbf{z}}(\underline{n})\} &= \sum_{k,\underline{k}} E\{\mathbf{g}(k)\,\mathbf{g}(\underline{k}+L)\}\, w(k)\, w(\underline{k}) \exp[-i2\pi\,(nk+\underline{n}\underline{k})/N] \\
&= \sum_{k} w(k)\, w(k-L) \exp[-i2\pi k/N - i2\pi \underline{n}(k-L)/N] \\
&= \exp(i2\pi \underline{n} L/N) \sum_{k} w(k)\, w(k-L) \exp[-i2\pi (n+\underline{n})k/N] \\
&= \exp(i2\pi \underline{n} L/N)\, W_1(n+\underline{n}) \equiv A(\underline{n}, n+\underline{n}).
\end{aligned} \tag{143}$$

Also, in a similar fashion,

$$\begin{aligned}
E\{\mathbf{z}(n)\,\tilde{\mathbf{z}}^*(\underline{n})\} &= \sum_{k} w(k)\, w(k-L) \exp[-i2\pi nk/N + i2\pi \underline{n}(k-L)/N] \\
&= \exp(-i2\pi \underline{n} L/N)\, W_1(n-\underline{n}) \equiv B(\underline{n}, n-\underline{n}).
\end{aligned} \tag{144}$$

Equations (143) and (144) contain complete covariance information about the complex spectral estimates $\{\mathbf{z}(n)\}$ and $\{\tilde{\mathbf{z}}(\underline{n})\}$ for arbitrary lag L. That is, temporal overlap is allowed and accounted for in these relations.

In terms of their real and imaginary components, let

$$\mathbf{z}(n) = \mathbf{x}(n) + i\,\mathbf{y}(n), \quad \tilde{\mathbf{z}}(\underline{n}) = \mathbf{u}(\underline{n}) + i\,\mathbf{v}(\underline{n}). \tag{145}$$

The means of all four real quantities are zero. Substitution of equation (145) into equations (143) and (144), and balancing of the real and imaginary parts, leads to the second-order relations:

$$
\begin{aligned}
E\{\mathbf{x}(n)\,\mathbf{u}(\underline{n})\} &= \frac{1}{2}\,B_r(\underline{n},n-\underline{n}) + \frac{1}{2}\,A_r(\underline{n},n+\underline{n}), \\
E\{\mathbf{y}(n)\,\mathbf{v}(\underline{n})\} &= \frac{1}{2}\,B_r(\underline{n},n-\underline{n}) - \frac{1}{2}\,A_r(\underline{n},n+\underline{n}), \\
E\{\mathbf{x}(n)\,\mathbf{v}(\underline{n})\} &= -\frac{1}{2}\,B_i(\underline{n},n-\underline{n}) + \frac{1}{2}\,A_i(\underline{n},n+\underline{n}), \\
E\{\mathbf{y}(n)\,\mathbf{u}(\underline{n})\} &= \frac{1}{2}\,B_i(\underline{n},n-\underline{n}) + \frac{1}{2}\,A_i(\underline{n},n+\underline{n}).
\end{aligned}
\tag{146}
$$

For this case, the last two expressions are *not* equal. In terms of the windows defined in equations (138), (143), and (144), there follows

$$
A(\underline{n},n) = \exp(i\,2\pi\,\underline{n}\,L/N)\,W_1(n), \quad B(\underline{n},n) = \exp(-i\,2\pi\,\underline{n}\,L/N)\,W_1(n). \tag{147}
$$

Equation (146) contains complete covariance information about the real and imaginary parts of complex spectral estimates $\{\mathbf{z}(n)\}$ and $\{\tilde{\mathbf{z}}(\underline{n})\}$, for arbitrary lag L.

7. SUMMARY

The first-order and second-order statistics for spectral estimates and correlation estimates obtained by means of weighted overlapped FFT processing have been obtained under a variety of conditions, including colored input data and wraparound in the time-delay domain. These results often require the definitions of new auxiliary functions that involve the weighting sequence $\{w(k)\}$, its length K, the amount of lag L between data segments, and the size N of the FFTs involved. In particular, a third-order correlation function of the weighting sequence was required to evaluate some of the covariances of interest. These auxiliary functions must then be applied in further summations to obtain the final closed-form results for the desired covariances.

APPENDIX A
INVERSE DISCRETE FOURIER TRANSFORM PROPERTIES

Function $C(k)$ is nonzero only for $|k| < K_c$. Define discrete Fourier transform

$$S_N(m) = \sum_k \exp(-i2\pi\, mk/N)\, C(k) \quad \text{for all } m, \tag{A-1}$$

where the sum is over all k. Function $S_N(m)$ has period N in m.

Now, define the inverse discrete Fourier transform

$$C_N(k) = \frac{1}{N} \sum_{m(N)} \exp(i2\pi\, km/N)\, S_N(m) \quad \text{for all } k, \tag{A-2}$$

where the sum is over any length N interval (one period) in m. Function $C_N(k)$ has period N in k; in fact,

$$C_N(k) = C(k) + C(k \pm N) + C(k \pm 2N) + \cdots. \tag{A-3}$$

That is, $C_N(k)$ is an aliased version of original function $C(k)$. However, if

$$N \geq 2K_c, \tag{A-4}$$

then none of the aliased lobes in equation (A-3) overlap the origin lobe $C(k)$. Then,

$$C(k) = \begin{Bmatrix} C_N(k) & \text{for } |k| < K_c \\ 0 & \text{for } |k| \geq K_c \end{Bmatrix} = G(k)\, C_N(k) \quad \text{for all } k, \tag{A-5}$$

where "gate function"

$$G(k) = \begin{Bmatrix} 1 & \text{for } |k| < N/2 \\ 0 & \text{for } |k| \geq N/2 \end{Bmatrix}. \tag{A-6}$$

Thus, if $N \geq 2K_c$, it is permissible to express

$$C(k) = G(k)\, \frac{1}{N} \sum_{m(N)} \exp(i2\pi\, km/N)\, S_N(m) \quad \text{for all } k. \tag{A-7}$$

The gate function $G(k)$ cannot be dropped in this relation.

As a special case, let $m(N) = 0 : N-1$ and define

$$S(m) = \begin{Bmatrix} S_N(m) & \text{for } m = 0 : N-1 \\ 0 & \text{otherwise} \end{Bmatrix}. \tag{A-8}$$

Then, for $N \geq 2K_c$, equation (A-7) becomes

$$C(k) = G(k)\,\frac{1}{N}\sum_m \exp(i2\pi\, km/N)\, S(m) \quad \text{for all } k. \tag{A-9}$$

Again, the gate function cannot be dropped in this relation.

APPENDIX B
APPROXIMATE EVALUATION OF $\theta(m,n,p)$

The function $\theta(m,n,p)$ was defined in equation (104) according to

$$\theta(m,n,p) = \sum_k w(k)\, w(k-m)\, w(k-n)\, w(k-p). \tag{B-1}$$

If weighting $\{w(k)\}$ varies slowly with k, as when sequence length K is large, the sum on k becomes approximately

$$\theta(m,n,p) \cong \int dx\, w(x)\, w(x-m)\, w(x-n)\, w(x-p). \tag{B-2}$$

For example, if Hann weighting

$$w(k) = \tfrac{1}{2} - \tfrac{1}{2}\cos(2\pi k/K) \ \text{ for } k = 1:K, \tag{B-3}$$

then

$$w(x) = \left\{ \begin{matrix} \tfrac{1}{2} - \tfrac{1}{2}\cos(2\pi x/K) \ \text{ for } \ 0 \le x \le K \\ 0 \ \text{ otherwise} \end{matrix} \right\}. \tag{B-4}$$

Although the integral in equation (B-2) can be carried out in closed form, it is extremely lengthy and would be time-consuming to calculate. An alternative approximation is to use

$$w(x) \cong \exp\left[-\frac{\pi^2}{K^2}\left(x - \frac{K}{2}\right)^2\right] \ \text{ for all } x, \tag{B-5}$$

which matches the Hann weighting (B-4) at its peak. Then, from equation (B-2), there follows

$$\begin{aligned} \theta(m,n,p) &\cong \frac{K}{2\sqrt{\pi}} \exp\left[-\frac{\pi^2}{4K^2}(3m^2 + 3n^2 + 3p^2 - 2mn - 2np - 2pm)\right] \\ &= \frac{K}{2\sqrt{\pi}} \exp\left[-\frac{\pi^2}{4K^2}\{m^2 + n^2 + p^2 + (m-n)^2 + (n-p)^2 + (p-m)^2\}\right]. \end{aligned} \tag{B-6}$$

This approximation yields

$$\theta(0,0,0) \cong \frac{K}{2\sqrt{\pi}} = 0.282\, K, \tag{B-7}$$

whereas the exact value is

$$\theta(0,0,0) = \sum_k w^4(k) = \frac{35}{128} K = 0.273\ K. \tag{B-8}$$

Also, the integral of equation (B-4) yields

$$\int dx\ w^4(x) = \frac{35}{128} K. \tag{B-9}$$

This suggests that the scale factor $1/(2\sqrt{\pi})$ in equation (B-6) be replaced by 35/128.

More generally, consider weighting function

$$w(x) = 1 - b - b\cos(2\pi\, x/K) \ \text{ for } 0 \le x \le K; \quad w(K/2) = 1. \tag{B-10}$$

For $b = 0.5$, $w(x)$ is Hann weighting, while for $b = 0.46$, $w(x)$ is Hamming weighting. There follows

$$w(x) \cong \exp\left[-b\frac{2\pi^2}{K^2}\left(x - \frac{K}{2}\right)^2\right] \ \text{ for all } x. \tag{B-11}$$

Then, equation (B-2) leads to

$$\theta(m,n,p) \cong \frac{K}{2\sqrt{2\pi\, b}} \exp\left[-b\frac{\pi^2}{2K^2}\{m^2 + n^2 + p^2 + (m-n)^2 + (n-p)^2 + (p-m)^2\}\right]. \tag{B-12}$$

In particular,

$$\theta(0,0,0) \cong \frac{K}{2\sqrt{2\pi\, b}} \quad (= 0.294\ K \ \text{ for } b = 0.46), \tag{B-13}$$

whereas the exact value is

$$\theta(0,0,0) = K\,(1 - 4b + 9b^2 - 10b^3 + 4.375b^4) \quad (= 0.287\ K \ \text{ for } b = 0.46). \tag{B-14}$$

These two functions of b in equations (B-13) and (B-14) are nearly equal for b in the neighborhood of (0.3, 0.5). In fact, they are equal at $b = 0.366948$.

INITIAL DISTRIBUTION LIST

Addressee	No. of Copies
Office of Naval Research (ONR 321US: H. South, J. Tague, D. Abraham)	3
University of Rhode Island (S. Kay)	1
University of Connecticut (P. Willett)	1
Defense Technical Information Center	2